EXAMPLE 6 Suppose that the average number of telephone calls arriving at the switchboard of a small corporation is 30 calls per hour. (i) What is the probability that no calls will arrive in a 3-minute period? (ii) What is the probability that more than five calls will arrive in a 5-minute interval? Assume that the number of calls arriving during any time period has a Poisson distribution. Assume that time is measured in minutes; then 30 calls per hour is equivalent to .5 calls per minute, so the *mean rate of occurrence* is .5 per minute. $P[\text{no calls in 3-minute period}] = e^{-\nu t} = e^{-(.5)(3)} = e^{-1.5} \approx .223$.

$$P[\text{more than five calls in 5-minute interval}] = \sum_{k=6}^{\infty} \frac{e^{-\nu t}(\nu t)^k}{k!}$$
$$= \sum_{k=6}^{\infty} \frac{e^{-(.5)(5)}(2.5)^k}{k!} \approx .042. \qquad ////$$

2.5 Geometric and Negative Binomial Distributions

Two other families of discrete distributions that play important roles in statistics are the geometric (or Pascal) and negative binomial distributions. The reason that we consider the two together is twofold; first, the geometric distribution is a special case of the negative binomial distribution, and, second, the sum of independent and identically distributed geometric random variables is negative binomially distributed, as we shall see in Chap. V. In Subsec. 3.3 of this chapter, the exponential and gamma distributions are defined. We shall see that in several respects the geometric and negative binomial distributions are discrete analogs of the exponential and gamma distributions.

Definition 6 Geometric distribution A random variable X is defined to have *geometric* (or *Pascal*) *distribution* if the density of X is given by

$$f_X(x) = f_X(x; p)$$
$$= \begin{Bmatrix} p(1-p)^x & \text{for } x = 0, 1, \ldots \\ 0 & \text{otherwise} \end{Bmatrix} = p(1-p)^x I_{\{0,1,\ldots\}}(x), \qquad (11)$$

where the parameter p satisfies $0 < p \le 1$. (Define $q = 1 - p$.) ////

Definition 7 Negative binomial distribution A random variable X with density

$$f_X(x) = f_X(x; r, p)$$
$$= \begin{cases} \binom{r+x-1}{x} p^r q^x = \binom{-r}{x} p^r (-q)^x & \text{for } x = 0, 1, 2, \ldots \\ 0 & \text{otherwise} \end{cases} \qquad (12)$$

where the parameters r and p satisfy $r = 1, 2, 3, \ldots$ and $0 < p \leq 1$ $(q = 1 - p)$, is defined to have a *negative binomial distribution*. The density given by Eq. (12) is called a *negative binomial density*.

////

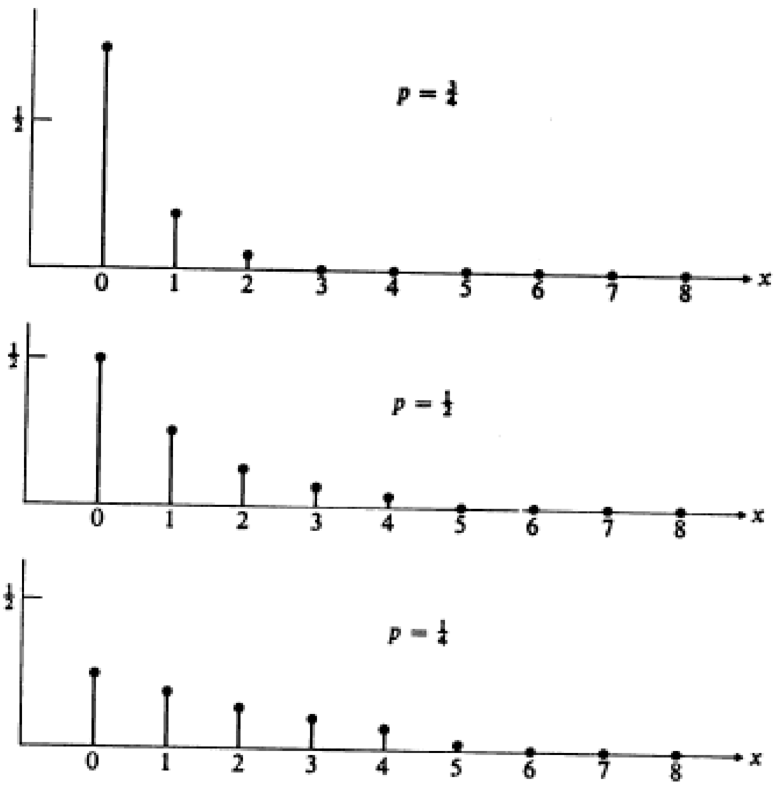

FIGURE 7
Geometric densities.

Theorem 9 If the random variable X has a geometric distribution, then

$$\mathscr{E}[X] = \frac{q}{p}, \qquad \operatorname{var}[X] = \frac{q}{p^2}, \qquad \text{and} \qquad m_X(t) = \frac{p}{1 - qe^t}. \tag{13}$$

PROOF Since a geometric distribution is a special case of a negative binomial distribution, Theorem 9 is a corollary of Theorem 11. ////

A random variable X that has a geometric distribution is often referred to as a discrete *waiting-time* random variable. It represents how long (in terms of the number of failures) one has to wait for a success.

Before leaving the geometric distribution, we note that some authors define the geometric distribution by assuming 1 (instead of 0) is the smallest mass point. The density then has the form

$$f(x; p) = p(1 - p)^{x-1} I_{\{1, 2, \ldots\}}(x), \tag{14}$$

and the mean is $1/p$, the variance is q/p^2, and the moment generating function is $pe^t/(1-qe^t)$.

Theorem 11 Let X have a negative binomial distribution; then

$$\mathscr{E}[X]=\frac{rq}{p}, \qquad \operatorname{var}[X]=\frac{rq}{p^2}, \qquad \text{and} \qquad m_X(t)=\left[\frac{p}{1-qe^t}\right]^r. \tag{15}$$

PROOF

$$m_X(t)=\mathscr{E}[e^{tX}]=\sum_{x=0}^{\infty}e^{tx}\binom{-r}{x}p^r(-q)^x$$

$$=\sum_{x=0}^{\infty}\binom{-r}{x}p^r(-qe^t)^x=\left[\frac{p}{1-qe^t}\right]^r$$

[see Eq. (33) in Appendix A].

$$m'_X(t)=p^r(-r)(1-qe^t)^{-r-1}(-qe^t)$$

and

$$m''_X(t)=rqp^r[q(r+1)e^{2t}(1-qe^t)^{-r-2}+e^t(1-qe^t)^{-r-1}];$$

hence

$$\mathscr{E}[X]=m'_X(t)\Big|_{t=0}=\frac{rq}{p}$$

and

$$\operatorname{var}[X]=m''_X(t)\Big|_{t=0}-(\mathscr{E}[X])^2=rqp^r[qp^{-r-2}(r+1)+p^{-r-1}]-\left(\frac{rq}{p}\right)^2$$

$$=\frac{rq^2}{p^2}+\frac{rq}{p}=\frac{rq}{p^2}.$$

3 CONTINUOUS DISTRIBUTIONS

In this section several parametric families of univariate probability density functions are presented. Sketches of some are included; the mean and variance (when they exist) of each are given.

3.1 Uniform or Rectangular Distribution

A very simple distribution for a continuous random variable is the uniform distribution. It is particularly useful in theoretical statistics because it is convenient to deal with mathematically.

Definition 10 **Uniform distribution** If the probability density function of a random variable X is given by

$$f_X(x) = f_X(x; a, b) = \frac{1}{b-a} I_{[a,b]}(x), \tag{21}$$

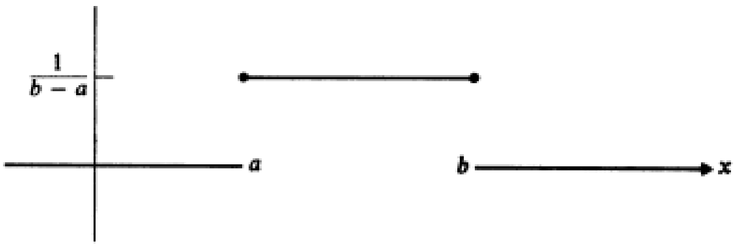

FIGURE 8
Uniform probability density.

where the parameters a and b satisfy $-\infty < a < b < \infty$, then the random variable X is defined to be *uniformly* distributed over the interval $[a, b]$, and the distribution given by Eq. (21) is called a *uniform distribution*.

////

Theorem 12 If X is uniformly distributed over $[a, b]$, then

$$\mathscr{E}[X] = \frac{a+b}{2}, \quad \text{var}[X] = \frac{(b-a)^2}{12}, \quad \text{and} \quad m_X(t) = \frac{e^{bt} - e^{at}}{(b-a)t}. \tag{22}$$

Theorem 12 If X is uniformly distributed over $[a, b]$, then

$$\mathscr{E}[X] = \frac{a+b}{2}, \qquad \text{var}[X] = \frac{(b-a)^2}{12}, \qquad \text{and} \qquad m_X(t) = \frac{e^{bt} - e^{at}}{(b-a)t}. \tag{22}$$

PROOF

$$\mathscr{E}[X] = \int_a^b x \frac{1}{b-a}\,dx = \frac{b^2 - a^2}{2(b-a)} = \frac{a+b}{2}.$$

$$\text{var}[X] = \mathscr{E}[X^2] - (\mathscr{E}[X])^2 = \int_a^b x^2 \frac{1}{b-a}\,dx - \left(\frac{a+b}{2}\right)^2$$

$$= \frac{b^3 - a^3}{3(b-a)} - \frac{(a+b)^2}{4} = \frac{(b-a)^2}{12}.$$

$$m_X(t) = \mathscr{E}[e^{tX}] = \int_a^b e^{tx} \frac{1}{b-a}\,dx = \frac{e^{bt} - e^{at}}{(b-a)t}.$$

////

The uniform distribution gets its name from the fact that its density is uniform, or constant, over the interval $[a, b]$. It is also called the *rectangular* distribution—the shape of the density is rectangular.

The cumulative distribution function of a uniform random variable is given by

$$F_X(x) = \left(\frac{x-a}{b-a}\right) I_{[a,b]}(x) + I_{(b,\infty)}(x). \tag{23}$$

3.2 Normal Distribution

A great many of the techniques used in applied statistics are based upon the normal distribution; it will frequently appear in the remainder of this book.

Definition 11 Normal distribution A random variable X is defined to be *normally* distributed if its density is given by

$$f_X(x) = f_X(x; \mu, \sigma) = \frac{1}{\sqrt{2\pi}\sigma} e^{-(x-\mu)^2/2\sigma^2}, \tag{24}$$

where the parameters μ and σ satisfy $-\infty < \mu < \infty$ and $\sigma > 0$. Any distribution defined by a density function given in Eq. (24) is called a *normal distribution*.

////

We have used the symbols μ and σ^2 to represent the parameters because these parameters turn out, as we shall see, to be the mean and variance, respectively, of the distribution.

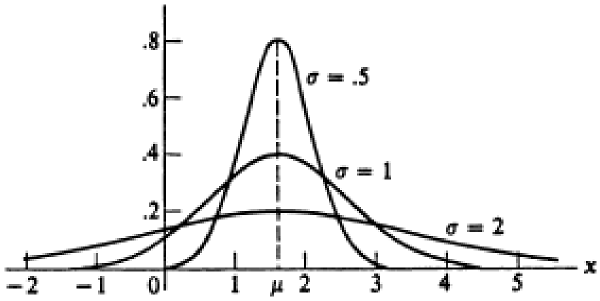

FIGURE 9
Normal densities.

One can readily check that the mode of a normal density occurs at $x = \mu$ and inflection points occur at $\mu - \sigma$ and $\mu + \sigma$. (See Fig. 9.) Since the normal distribution occurs so frequently in later chapters, special notation is introduced for it. If random variable X is normally distributed with mean μ and variance σ^2, we will write $X \sim N(\mu, \sigma^2)$. We will also use the notation $\phi_{\mu,\sigma^2}(x)$ for the density of $X \sim N(\mu, \sigma^2)$ and $\Phi_{\mu,\sigma^2}(x)$ for the cumulative distribution function.

If the normal random variable has mean 0 and variance 1, it is called a *standard* or *normalized* normal random variable. For a standard normal random variable the subscripts of the density and distribution function notations are dropped; that is,

$$\phi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} \qquad \text{and} \qquad \Phi(x) = \int_{-\infty}^{x} \phi(u)\, du. \tag{25}$$

Since $\phi_{\mu,\sigma^2}(x)$ is given to be a density function, it is implied that

$$\int_{-\infty}^{\infty} \phi_{\mu,\sigma^2}(x)\, dx = 1,$$

Theorem 13 If X is a normal random variable,

$$\mathscr{E}[X] = \mu, \qquad \operatorname{var}[X] = \sigma^2, \qquad \text{and} \qquad m_X(t) = e^{\mu t + \sigma^2 t^2/2}. \tag{26}$$

PROOF

$$\begin{aligned} m_X(t) = \mathscr{E}[e^{tX}] &= e^{t\mu}\mathscr{E}[e^{t(X-\mu)}] \\ &= e^{t\mu} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{t(x-\mu)} e^{-(1/2\sigma^2)(x-\mu)^2}\, dx \\ &= e^{t\mu} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-(1/2\sigma^2)[(x-\mu)^2 - 2\sigma^2 t(x-\mu)]}\, dx. \end{aligned}$$

If we complete the square inside the bracket, it becomes

$$\begin{aligned} (x-\mu)^2 - 2\sigma^2 t(x-\mu) &= (x-\mu)^2 - 2\sigma^2 t(x-\mu) + \sigma^4 t^2 - \sigma^4 t^2 \\ &= (x - \mu - \sigma^2 t)^2 - \sigma^4 t^2, \end{aligned}$$

$$m_X(t) = e^{t\mu} e^{\sigma^2 t^2/2} \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{\infty} e^{-(x-\mu-\sigma^2 t)^2/2\sigma^2} \, dx.$$

The integral together with the factor $1/\sqrt{2\pi}\sigma$ is necessarily 1 since it is the area under a normal distribution with mean $\mu + \sigma^2 t$ and variance σ^2. Hence,

$$m_X(t) = e^{\mu t + \sigma^2 t^2/2}.$$

On differentiating $m_X(t)$ twice and substituting $t = 0$, we find

$$\mathscr{E}[X] = m_X'(0) = \mu$$

and

$$\text{var}[X] = \mathscr{E}[X^2] - (\mathscr{E}[X])^2 = m_X''(0) - \mu^2 = \sigma^2,$$

thus justifying our use of the symbols μ and σ^2 for the parameters. ////

Since the indefinite integral of $\phi_{\mu,\sigma^2}(x)$ does not have a simple functional form, one can only exhibit the cumulative distribution function as

$$\Phi_{\mu,\sigma^2}(x) = \int_{-\infty}^{x} \phi_{\mu,\sigma^2}(u) \, du. \qquad (27)$$

Theorem 14 If $X \sim N(\mu, \sigma^2)$, then

$$P[a < X < b] = \Phi\left(\frac{b-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right). \qquad (28)$$

Remark $\Phi(x) = 1 - \Phi(-x)$.

EXAMPLE 13 Suppose that an instructor assumes that a student's final score is the value of a normally distributed random variable. If the instructor decides to award a grade of A to those students whose score exceeds $\mu + \sigma$, a B to those students whose score falls between μ and $\mu + \sigma$, a C if a score falls between $\mu - \sigma$ and μ, a D if a score falls between $\mu - 2\sigma$ and $\mu - \sigma$, and an F if the score falls below $\mu - 2\sigma$, then the proportions of each grade given can be calculated. For example, since

$$P[X > \mu + \sigma] = 1 - P[X < \mu + \sigma] = 1 - \Phi\left(\frac{\mu + \sigma - \mu}{\sigma}\right)$$
$$= 1 - \Phi(1) \approx .1587,$$

one would expect 15.87 percent of the students to receive A's. ////

$$m_X(t) = e^{t\mu}e^{\sigma^2t^2/2}\frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{\infty} e^{-(x-\mu-\sigma^2t)^2/2\sigma^2}\,dx.$$

The integral together with the factor $1/\sqrt{2\pi}\sigma$ is necessarily 1 since it is the area under a normal distribution with mean $\mu + \sigma^2 t$ and variance σ^2. Hence,

$$m_X(t) = e^{\mu t + \sigma^2 t^2/2}.$$

On differentiating $m_X(t)$ twice and substituting $t = 0$, we find

$$\mathscr{E}[X] = m_X'(0) = \mu$$

and

$$\text{var}\,[X] = \mathscr{E}[X^2] - (\mathscr{E}[X])^2 = m_X''(0) - \mu^2 = \sigma^2,$$

thus justifying our use of the symbols μ and σ^2 for the parameters. ////

Since the indefinite integral of $\phi_{\mu,\sigma^2}(x)$ does not have a simple functional form, one can only exhibit the cumulative distribution function as

$$\Phi_{\mu,\sigma^2}(x) = \int_{-\infty}^{x} \phi_{\mu,\sigma^2}(u)\,du. \qquad (27)$$

Theorem 14 If $X \sim N(\mu, \sigma^2)$, then

$$P[a < X < b] = \Phi\left(\frac{b-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right). \qquad (28)$$

Remark $\Phi(x) = 1 - \Phi(-x)$.

EXAMPLE 13 Suppose that an instructor assumes that a student's final score is the value of a normally distributed random variable. If the instructor decides to award a grade of A to those students whose score exceeds $\mu + \sigma$, a B to those students whose score falls between μ and $\mu + \sigma$, a C if a score falls between $\mu - \sigma$ and μ, a D if a score falls between $\mu - 2\sigma$ and $\mu - \sigma$, and an F if the score falls below $\mu - 2\sigma$, then the proportions of each grade given can be calculated. For example, since

$$P[X > \mu + \sigma] = 1 - P[X < \mu + \sigma] = 1 - \Phi\left(\frac{\mu + \sigma - \mu}{\sigma}\right)$$
$$= 1 - \Phi(1) \approx .1587,$$

one would expect 15.87 percent of the students to receive A's. ////

EXAMPLE 14 Suppose that the diameters of shafts manufactured by a certain machine are normal random variables with mean 10 centimeters and standard deviation .1 centimeter. If for a given application the shaft must meet the requirement that its diameter fall between 9.9 and 10.2 centimeters, what proportion of the shafts made by this machine will meet the requirement?

$$P[9.9 < X < 10.2] = \Phi\left(\frac{10.2 - 10}{.1}\right) - \Phi\left(\frac{9.9 - 10}{.1}\right)$$
$$= \Phi(2) - \Phi(-1) \approx .9772 - .1587 = .8185. \qquad ////$$

3.3 Exponential and Gamma Distributions

Two other families of distributions that play important roles in statistics are the (negative) exponential and gamma distributions, which are defined in this subsection. The reason that the two are considered together is twofold; first, the exponential is a special case of the gamma, and, second, the sum of independent identically distributed exponential random variables is gamma-distributed, as we shall see in Chap. V.

Definition 12 Exponential distribution If a random variable X has a density given by

$$f_X(x; \lambda) = \lambda e^{-\lambda x} I_{[0, \infty)}(x), \qquad (29)$$

where $\lambda > 0$, then X is defined to have an (negative) *exponential distribution.* ////

Definition 13 Gamma distribution If a random variable X has density given by

$$f_X(x; r, \lambda) = \frac{\lambda}{\Gamma(r)} (\lambda x)^{r-1} e^{-\lambda x} I_{[0, \infty)}(x), \qquad (30)$$

where $r > 0$ and $\lambda > 0$, then X is defined to have a *gamma distribution.* $\Gamma(\cdot)$ is the gamma function and it is discussed in Appendix A. ////

Remark If in the gamma density $r = 1$, the gamma density specializes to the exponential density. ////

Theorem 15 If X has an exponential distribution, then

$$\mathscr{E}[X] = \frac{1}{\lambda}, \qquad \text{var}[X] = \frac{1}{\lambda^2}, \qquad \text{and} \qquad m_X(t) = \frac{\lambda}{\lambda - t} \qquad \text{for} \qquad t < \lambda. \qquad (31)$$

Theorem 16 If X has a gamma distribution with parameters r and λ, then

$$\mathscr{E}[X] = \frac{r}{\lambda}, \qquad \text{var}[X] = \frac{r}{\lambda^2}, \qquad \text{and} \qquad m_X(t) = \left(\frac{\lambda}{\lambda - t}\right)^r \qquad \text{for } t < \lambda. \tag{32}$$

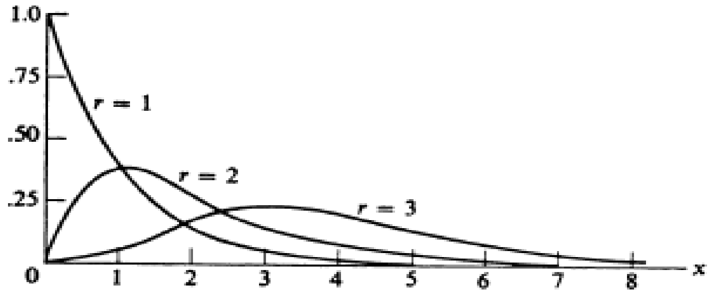

FIGURE 11
Gamma densities ($\lambda = 1$).

PROOF

$$\begin{aligned} m_X(t) &= \mathscr{E}[e^{tX}] \\ &= \int_0^\infty \frac{\lambda^r}{\Gamma(r)} e^{tx} x^{r-1} e^{-\lambda x}\, dx \\ &= \left(\frac{\lambda}{\lambda - t}\right)^r \int_0^\infty \frac{(\lambda - t)^r}{\Gamma(r)} x^{r-1} e^{-(\lambda - t)x}\, dx = \left(\frac{\lambda}{\lambda - t}\right)^r. \\ m_X'(t) &= r\lambda^r(\lambda - t)^{-r-1} \end{aligned}$$

and

$$m_X''(t) = r(r+1)\lambda^r(\lambda - t)^{-r-2};$$

hence

$$\mathscr{E}[X] = m_X'(0) = \frac{r}{\lambda}$$

and

$$\begin{aligned} \text{var}[X] &= \mathscr{E}[X^2] - (\mathscr{E}[X])^2 \\ &= m_X''(0) - \left(\frac{r}{\lambda}\right)^2 = \frac{r(r+1)}{\lambda^2} - \left(\frac{r}{\lambda}\right)^2 = \frac{r}{\lambda^2}. \qquad //// \end{aligned}$$

The exponential distribution has been used as a model for lifetimes of various things. When we introduced the Poisson distribution, we spoke of certain happenings, for example, particle emissions, occurring in time. The length of the time interval between successive happenings can be shown to have an exponential distribution provided that the number of happenings in a fixed